

2006

Using Self-Organizing Maps to Recognize Acoustic Units Associated with Information Content in Animal Vocalizations

John Placer

Northern Arizona University

C. N. Slobodchikoff

Northern Arizona University

Jason Burns

Northern Arizona University

Jeffrey Placer

Northern Arizona University

Ryan Middleton

Northern Arizona University

Follow this and additional works at: http://animalstudiesrepository.org/acwp_vsm



Part of the [Animals Commons](#), [Animal Studies Commons](#), and the [Other Animal Sciences Commons](#)

Recommended Citation

Placer, J., Slobodchikoff, C. N., Burns, J., Placer, J., & Middleton, R. (2006). Using self-organizing maps to recognize acoustic units associated with information content in animal vocalizations. *The Journal of the Acoustical Society of America*, 119(5), 3140-3146.

This Article is brought to you for free and open access by the Articles, Studies, and Book Chapters at Animal Studies Repository. It has been accepted for inclusion in Animal Science, Veterinary Medicine, and Zoology by an authorized administrator of Animal Studies Repository. For more information, please contact eyahner@humanesociety.org.

Using Self-Organizing Maps to Recognize Acoustic Units Associated with Information Content in Animal Vocalizations

John Placer, C. N. Slobodchikoff, Jason Burns, Jeffrey Placer, and Ryan Middleton
Northern Arizona University

ABSTRACT

Kohonen self-organizing neural networks, also called self-organizing maps (SOMs), have been used successfully to recognize human phonemes and in this way to aid in human speech recognition. This paper describes how SOMs also can be used to associate specific information content with animal vocalizations. A SOM was used to identify acoustic units in Gunnison's prairie dog alarm calls that were vocalized in the presence of three different predator species. Some of these acoustic units and their combinations were found exclusively in the alarm calls associated with a particular predator species and were used to associate predator species information with individual alarm calls. This methodology allowed individual alarm calls to be classified by predator species with an average of 91% accuracy. Furthermore, the topological structure of the SOM used in these experiments provided additional insights about the acoustic units and their combinations that were used to classify the target alarm calls. An important benefit of the methodology developed in this paper is that it could be used to search for groups of sounds associated with information content for any animal whose vocalizations are composed of multiple simultaneous frequency components.

1. INTRODUCTION

Research on animal vocalizations that contain multiple simultaneous frequency components often focuses on the general, aggregate structure of those vocalizations. Frequency and time parameters are gathered and an entire vocalization is characterized as a collection of frequency-time attributes. In some studies these aggregate properties of sounds have been compared for overlap using spectrogram cross correlation (Clark *et al.*, 1987). Alternatively information has been extracted and analyzed statistically using multivariate methods such as Principal Components Analysis or Discriminant Function Analysis (Eakle *et al.*, 1989; Galeotti and Pavan, 1991; Slobodchikoff *et al.*, 1991; Galeotti *et al.*, 1993; Blackshaw *et al.*, 1996; Hill and Lill, 1998; Slobodchikoff *et al.*, 1991; Ackers and Slobodchikoff, 1999; Lengagne, 2001; Sousa-Lima *et al.*, 2002; Baker and Logue, 2003)

As an example of studies that focus on the general aggregate structure of vocalizations consider Fig. 1, which shows how sound spectrograms can be used to visually distinguish between different types of alarm calls vocalized by Gunnison's prairie dogs. The shape of the spectrogram on the left is typical of

dog-elicited alarm calls and the shape of the spectrogram on the right is typical of coyote-elicited alarm calls. In past studies analyses related to the general structure of alarm calls have been performed on data by measuring various macroscopic properties of collections of similar sound spectrograms of alarm calls. Properties such as ascending slope, descending slope, dominant frequency, and subdominant frequency were measured and the data was then subjected to discriminant function analysis (Slobodchikoff *et al.*, 1991). Note that the sound spectrograms in Fig. 1 were created from the same database that provided the data for the experiments discussed in this paper.

In the types of studies referred to above, the macroscopic structure of vocalizations is examined while the detailed internal structure, the acoustic microstructure, of the vocalizations is left unexamined and unappreciated. One striking illustration of this is the case of repetitive vocalizations of animal signals; it is generally assumed that repeated vocalizations from a given animal convey the same information even when the frequency-time parameters of their sound spectrograms vary. However, alarm calls are an example of repetitive vocalizations whose internal variations might be highly significant. Alarm calls are often given as a series or *bout* of repetitive elements, as in the case of alarm calls vocalized by Gunnison's prairie dog (*Cynomys gunnisoni*) in the presence of predators (Slobodchikoff *et al.*, 1991). This paper presents a methodology for investigating the possibility that the variations in the acoustic properties of animal alarm calls are highly significant because these variations indicate the existence of distinct acoustic structures that encode information.

A link between acoustic microstructures and information encoding in alarm calls has been suggested by earlier studies of Gunnison's prairie dog vocalizations. In two studies (Placer and Slobodchikoff, 2000, 2001), composites of frequency ratio information taken from small partitions of Gunnison's prairie dog alarm calls were used to classify those calls according to the species of predator that was present when the calls were vocalized. These results suggested that information about predator species was encoded in the acoustic microstructure of the alarm calls examined. However, the composite frequency ratio distributions did not themselves reveal the details of the acoustic structure of individual alarm call partitions nor did they provide any clues about how these small-scale acoustics might be organized to encode information in the macroscopic structure of an alarm call.

In a later study (Placer and Slobodchikoff, 2004), Gunnison's prairie dog alarm calls were decomposed into numerous equal-sized partitions and the four most energetic frequency components of each partition were extracted. The two lowest frequencies of the four were then used to identify the acoustic unit associated with any given partition. When the acoustic units created by this process were examined, it was discovered that there were unique acoustic units and combinations of acoustic units that occurred exclusively in vocalizations associated with specific predator species. These exclusive sounds were used to classify alarm calls according to the species of predator that was present when the calls were vocalized. However, the accuracy of the classifications of individual alarm calls were modest and varied significantly according to different criteria that were applied in the study. Nonetheless, the results of this research also suggested a potential link between acoustic microstructure and information encoding and they began to reveal some of the important features of the acoustic microstructure of alarm calls.

In the work described in this paper new, more accurate techniques are developed that are used to recognize small-scale acoustic structures in animal alarm calls that might be used to encode information. These techniques are based on the self-organizing map (SOM) algorithm of Kohonen (1998). A SOM or Kohonen map is an artificial neural network that classifies input data into a number of clusters that are based on the features of the input data itself. In other words, a SOM is not told what the categories of interest are; the SOM uses its input data to discover for itself what the relevant categories are. This is a particularly important attribute for the study discussed in this paper where the relevant categories are acoustic units whose structures or features are not yet well understood.

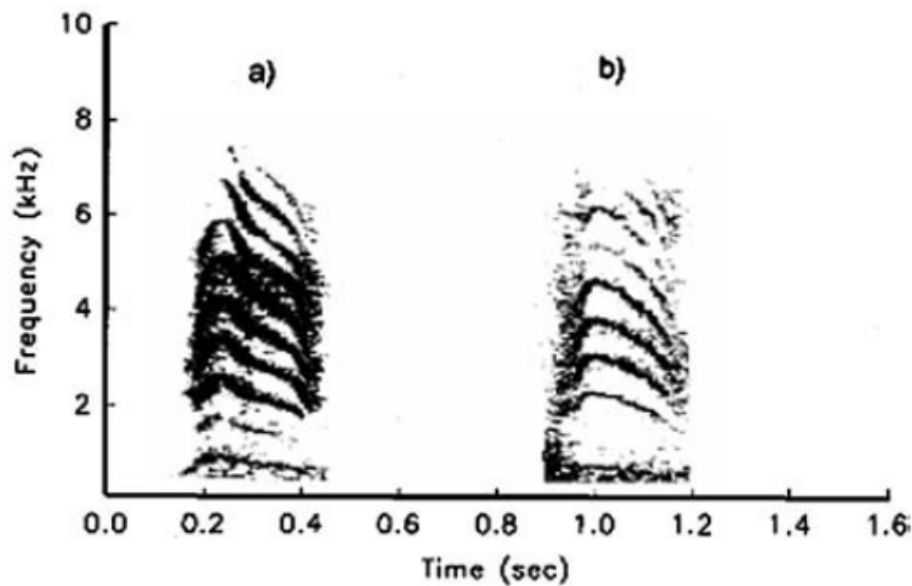


FIG. 1. Sound spectrograms showing a typical dog-elicited alarm call on the left-hand side and a typical coyote-elicited alarm call on the right-hand side.

Self-organizing maps have been shown to be effective in the automatic recognition of human speech. Kohonen used a SOM to create a recognition system for human speech that could detect phonemes from a continuous speech signal (Kohonen, 1988). This “phonetic typewriter,” as Kohonen called it, was used to recognize both Finnish and Japanese phonemes. Since then SOMs have been used for the phoneme-based recognition of a number of other languages including Chinese (Wu *et al.*, 1992) and Slovene (Mihelic *et al.*, 1992). SOMs have also been used as more general feature extractors of human speech that can utilize contextual information (Kangas *et al.*, 1992). In sender recognition experiments with six mouse lemur males, Zimmerman and Lerch used a high performance Kohonen map trained for continuous human speech recognition. The success rate in recognizing these six individual males varied from 55.6% to 100% (Zimmerman and Lerch, 1993).

This paper describes how SOMs can be used to characterize and identify classes of acoustic units that are associated with specific information content in animal vocalizations. A SOM was trained to identify clusters of acoustic units in Gunnison's prairie dog alarms where each cluster contained sounds with similar acoustic properties. Individual sounds belonging to specific clusters as well as combinations of these sounds were found to be associated exclusively with alarm calls vocalized in the presence of a specific predator species. These sounds and sound combinations were used with high accuracy to classify the targeted alarm calls according to the species of predator that was present when the alarm calls were vocalized. Furthermore, the topological structure of the output layer of the SOM was shown to reveal interesting characteristics of these sounds and their combinations. The details of these experiments are presented in the following sections of this paper.

2. MATERIALS AND METHODS

This study utilized digitized versions of Gunnison's prairie dog alarm calls that were recorded over a period of 10 years (1988–1997) at two separate prairie dog colonies. Uher Model No. CR160AV and

Sony TC-D5PRO II cassette recorders and a Sennheiser ME-88 shotgun microphone were used to record the alarm calls. Additional details about this library of alarm calls can be found in Placer and Slobodchikoff (2004). A total of 300 alarm calls was selected from this database where 100 alarm calls were vocalized in the presence of dogs, 100 in the presence of coyotes, and 100 in the presence of hawks. A system of computer programs was created by one of the authors (J.P.) that performed all the data analysis presented in this paper. Some of these programs were implemented using MATLAB, a high-performance numeric computation software package, and other programs were implemented using the computer language Java.

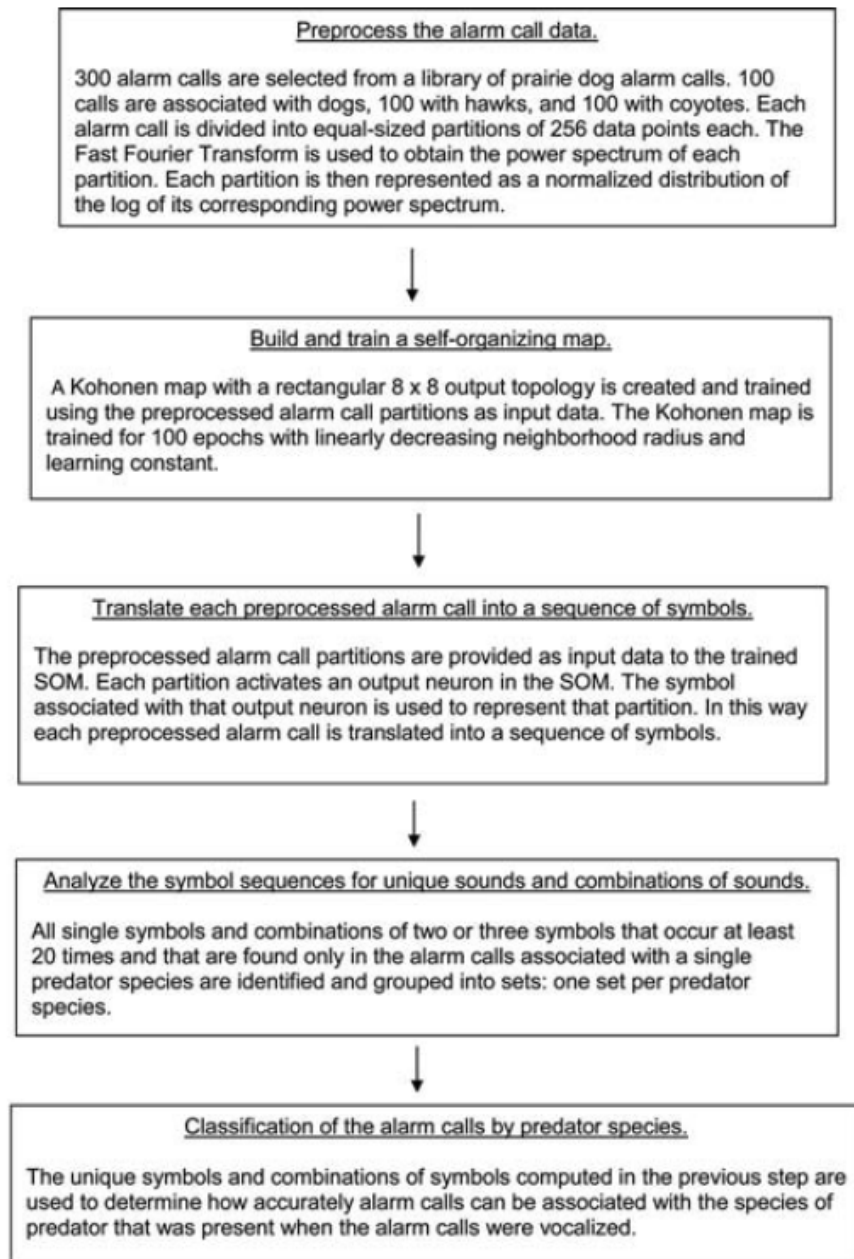


FIG. 2. The experimental protocol used to classify prairie dog alarm calls by the predator species that was present when the calls were vocalized.

The experiments discussed in this paper consisted of five distinct steps. First, each of the 300 alarm calls utilized in the experiments was preprocessed using the fast Fourier transform. Second, a self-organizing map was trained using the preprocessed data as input. Third, the trained SOM was used to translate the preprocessed alarm call data into sequences of symbols where each distinct symbol represented a distinct cluster of similar acoustic units. Fourth, the sequences of symbols were examined for single symbols and combinations of two or three symbols that were found exclusively in the alarm calls associated with a specific predator species. Fifth, these unique symbols were then used to classify the alarm calls according to the predator species present when the calls were vocalized. Furthermore, relationships among the unique symbols were examined. A graphical representation of these steps is given in Fig. 2. Each of these steps is described in detail in the sections provided below.

Preprocess the alarm call data

Each of the 300 alarm calls utilized in this study was preprocessed. Every alarm call was divided into equal-sized partitions of 256 data points each. The fast Fourier transform (FFT) was then applied to each partition's data points in order to compute a power spectrum of the frequencies contained in that partition. The alarm call data had been recorded at 44 100 samples per second and each partition contained 256 data points so the resolution of the FFT was just over 172 Hertz ($44100/256$). Since the upper limit of the frequencies that needed to be examined was only 9000 Hertz, each power spectrum contained 52 power values (the floor of $9000/172$). A normalized distribution of the log of each power spectrum was then used to represent each partition. When the preprocessing just described was complete, each alarm call had been translated into a sequence of 25 vectors, where each vector contained 52 values.

Build and train a self-organizing map

A Kohonen map (SOM) was implemented for this experiment using the language Java. A SOM has two layers: an input layer and an output layer. Each neuron in the input layer is connected to each neuron in the output layer where each connection between an input neuron and an output neuron contains a weight. The input layer simply holds the values of the current input vector. The Kohonen map used in the experiments discussed in this paper had 52 input neurons, one neuron for each value in the normalized distribution of the log of the power spectrum of an alarm call partition. Since each input neuron was connected to each output neuron, every output neuron had 52 weights associated with it.

The output neurons of a SOM can be arranged in a number of different spatial or topological structures. The SOM used in these experiments had an output layer that consisted of a simple 8×8 rectangular grid of neurons. During training, when an input vector is provided to a SOM, the output neuron whose weight vector is closest to the value of the input vector (i.e., the winning neuron) has its weights adjusted so that they more closely resemble the values of the input vector. Furthermore, not only are the weights of the winning output neuron adjusted, but the neurons in a neighborhood surrounding that output neuron also have their weights adjusted to more closely resemble the input vector. The closeness between two vectors was determined by a simple Euclidean distance measure.

The neighborhood radius for the SOM used in this experiment began at 4 neurons and was decreased linearly as training progressed. The degree of adjustment of weights in a SOM is determined by a learning constant. When the weight vector on the winning output neuron is adjusted to align more closely with the current input vector, the learning constant limits the amount of this alignment. With a learning constant of 1.0, the weight vector would be moved so that it is exactly aligned with the input vector. A learning constant of 0.5 would cause the weight vector to be moved half of the distance to complete alignment. In this experiment the learning constant began with the value 0.7 and was decreased linearly to the value 0.05 as training progressed. Both the neighborhood radius and the learning constant reached

their final values of 0 and 0.05, respectively, one-half of the way through the SOM training session. In the second half of the session, called the convergence phase, the neighborhood vanishes so that only the winning neuron itself is adjusted with each input. Furthermore, the learning constant is quite small so that the SOM slowly refines and adjusts the weights on its winning output neurons.

TABLE 1. The mapping of the Kohonen map's output neurons to symbol names. Each cell represents an output neuron and the text in each cell is the name of the symbol associated with that cell.

S1	S2	S3	S4	S5	S6	S7	S8
S9	S10	S11	S12	S13	S14	S15	S16
S17	S18	S19	S20	S21	S22	S23	S24
S25	S26	S27	S28	S29	S30	S31	S32
S33	S34	S35	S36	S37	S38	S39	S40
S41	S42	S43	S44	S45	S46	S47	S48
S49	S50	S51	S52	S53	S54	S55	S56
S57	S58	S59	S60	S61	S62	S63	S64

TABLE 2. Statistics for the alarm calls correctly classified by predator species with the standard deviations for the final averages given as parenthesized values.

Run #	Dog %Correct	Hawk %Correct	Coyote %Correct	Average %Correct
1	82	100	91	91
2	90	100	88	92.67
3	75	100	91	88.67
4	91	100	95	95.33
5	95	100	91	95.33
6	81	100	90	90.33
7	92	100	83	91.67
8	82	100	88	90.00
9	92	100	89	93.67
10	81	100	83	88.00
Average %Correct	86.1 (6.64)	100 (0)	88.9 (3.70)	91.67 (2.57)

A training session of a SOM consists of repeatedly presenting the vectors in the input data set (the preprocessed alarm calls in our case) to the SOM and allowing the SOM to adjust the weights on its connections. The SOM used in this experiment was trained for 100 epochs. One epoch is completed when all of the vectors in the input data set have been presented to the SOM. One major result of training a SOM is that the weights of its output neurons approach values that represent groups or clusters of similar vectors in the input data set. In effect, each output neuron becomes an exemplar of a class of similar vectors in the input data set. Recall that each vector in the input data set used in this experiment represents the acoustic unit found in a partition of one of the 300 target alarm calls. Therefore, the output nodes of the SOM used in this experiment represented clusters of similar acoustic units found in the partitions of the alarm calls utilized for the experiment.

Another major result of training a SOM is found in the topology or spatial ordering of its output layer. Since neighborhoods of output neurons have their weights adjusted together, similar input patterns will be recognized by output neurons that are spatially close to each other in the output layer of the SOM. In the

experiments discussed here, this means that similar acoustic units will be found in neurons that lie close to each other in the 8×8 output grid of the SOM.

Translate each preprocessed alarm call into a sequence of symbols

After the SOM was trained in this experiment, it was then used to translate each preprocessed alarm call into a sequence of symbols. To understand how this was done consider that the output layer of the SOM was an 8×8 matrix of neurons. Each neuron had a unique symbol associated with it according to its position in the matrix. The mapping of symbols to output neurons is shown in Table 1. Each cell in the matrix in Table 1 represents an output neuron in the SOM used in these experiments. The text contained in any given cell is the name of the symbol associated with that cell. After training the SOM, the preprocessed partitions composing each alarm call were once again provided as input to the trained SOM. Each input vector (i.e., preprocessed partition) provided to the SOM was replaced by the symbol associated with the winning neuron produced by that input vector; thus, each partition in each alarm call was replaced by a specific symbol. In this way, each preprocessed alarm call was translated into a sequence of symbols. For example, the sequence of symbols

S9S46S40S40S23S35S35S35S42S35S23S35S42S35S35S35S8S40

was produced when the trained SOM was used to translate one of the preprocessed alarm calls that had been vocalized in the presence of a dog.

Analyze the symbol sequences for unique sounds and combinations of sounds

After this translation process in which each preprocessed alarm call was expressed as a sequence of symbols, the symbols contained in the translated alarm calls were analyzed in order to identify any single symbols (monograms), neighboring pairs of symbols (digrams) and sequences of three symbols (trigrams) that were found only in the alarm calls associated with a particular species of predator. During this part of the analysis only monograms, digrams, and trigrams that occurred at least 20 times in the target alarm calls were examined. This precaution was meant to guard against the use of acoustic units that rarely occurred and that might be artifacts of the arbitrary way in which alarm calls were partitioned. Three sets of symbols were constructed: one for monograms, digrams, and trigrams that occurred only in alarm calls associated with coyotes, another set for monograms, digrams, and trigrams that occurred only in alarm calls associated with dogs, and a third set for monograms, digrams, and trigrams that occurred only in alarm calls associated with hawks.

Classification of the alarm calls by predator species

In principle, any alarm call that could be shown to contain a monogram, digram, or trigram that was uniquely associated with a given predator species could itself be unambiguously associated with that same predator species. Thus, the three sets of unique monograms, digrams, and trigrams described above were used to classify the alarm calls by predator species. The percentage of alarm calls that could be identified unambiguously in this way for a given predator species was considered the percentage of correctly identified alarm calls for that predator species.

3. RESULTS

Kohonen maps are initialized with random connection weights before training. One consequence of this is that the final connection weights that result from any given training session will vary. This, in turn, means that the results of translating alarm calls into symbol sequences and the associated classification accuracies will vary among a number of separately trained SOMs. Therefore, in order to determine

average classification accuracies, 10 experimental runs, each performed with a separately trained Kohonen map (SOM), were executed. The data resulting from these 10 experimental runs are given in Table 2.

Each row in Table 2 except the last gives the data associated with a separate experimental run. The last row provides the means of all the data generated in the 10 experiments with standard deviation values given in parentheses. The first value of each row simply gives an identifying number for that experimental run. The second, third, and fourth values give the percentage of alarm calls classified correctly for the predator species dog, hawk, and coyote, respectively. The last value provides the average correct classification for that run across all species of predators.

Table 1, which maps symbols to output neurons, is expanded in Table 3 to include information about the monograms, digrams, and trigrams that were found to be unique to a particular predator species. The name in bold at the top center of each cell (output neuron) is the name of the symbol associated with that cell. If that symbol is found at the start of one or more monograms, digrams or trigrams found exclusively in alarm calls associated with only one predator species, then all of those monograms, digrams, and trigrams are listed below the symbol. Cells that contain such lists also specify the relevant predator species in parentheses next to the cell's symbol name at the top of the cell. Symbols associated with cells that contain no lists below their symbol name simply do not occur at the beginning of a monogram, digram or trigram that is unique to a particular predator species.

To better explain Table III some examples are needed. Consider the cell named **S24**. This cell contains a list with only one digram (S24S21), which begins with symbol S24. This digram is found only in the alarm calls associated with the species coyote; the species name is shown in parentheses next to the cell's symbol name. Now consider cell **S62**. The list in this cell contains two digrams, S62S62 and S62S37, both of which begin with symbol S62. The digram S62S62 indicates that the sound associated with symbol S62 is held across two partitions of an alarm call. Notice again that the predator species coyote is specified in parentheses next to the cell's symbol name. Finally, consider cell **S4**. The predator species specified in this cell is dog and the list provided in the cell simply contains S4. This means that the monogram S4 occurs only in alarm calls associated with dogs; therefore, no digrams or trigrams associated exclusively with dogs begin with the symbol S4.

Other types of information can also be gained from Table 3. For example, notice that the cells associated with dogs and coyotes each tend to have short lists of unique monograms, digrams, and trigrams. However, the cells associated with hawks tend to contain long lists. A closer look at these long lists reveals that combinations of a small group of symbols (S35, S51, S53, S58, and S60) are used to create most of the digrams and trigrams in all of the lists associated with hawks. These same symbols are used in differing orders and combinations throughout the lists.

4. DISCUSSION

Kohonen maps were used to associate individual alarm calls with specific predator species with average accuracies of greater than 91%; these results demonstrate that Kohonen maps are useful tools for exploring and identifying sounds associated with specific information content in animal vocalizations. Furthermore, analysis of the digrams and trigrams specific to a particular predator species revealed certain types of symbol patterns associated with specific predator species. For example, in Table 3, digrams and trigrams associated with hawks were regrouped into new lists where the same two symbols were used in each digram or trigram. This regrouping revealed how particular symbols can occur in different orders or be expressed for varying lengths of time or both. As an example, consider the list

below that is composed of digrams and trigrams associated only with hawks, where each digram and trigram contains only the symbols S53 and S60:

S53S60, S53S53S60, S60S53, S60S53S53.

The symbol sequences in the list above indicate that the symbol S53 can occur either before or after the symbol S60. In other words, the relative order of the two sounds is not fixed. Furthermore, recall that S53 represents a sound (acoustic unit) expressed in a single partition of an alarm call whereas S53S53 represents the sound S53 expressed across two partitions of an alarm call. Thus, in the list above the sound represented by symbol S53 can be expressed for a single partition of time (S53S60, S60S53) or it can be held across two partitions of time (S53S53S60, S60S53S53). Similar kinds of relationships can be found in digrams and trigrams containing different pairs of symbols. Whether these patterns of symbols represent a first glimpse of lexical structures or a first glimpse of syntactic structures in prairie dog alarm calls is yet to be determined. However, the patterns are clearly evident and they are seen to occur in vocalizations restricted to a specific predator species.

It should also be mentioned that the variety of symbols (sounds) associated with specific predator species in Table 3 is probably less diverse than it might seem at first. Symbols that lie close to each other in the Kohonen map output layer and that are unique to the same predator species might well represent slight variations of the same basic sound. For example, consider cells **S25** and **S26**, which list digrams unique to alarm calls associated with dogs. These cells are located next to each other and the digram in each cell only contains the symbol named by that cell. Thus, it is likely that symbols S25 and S26 refer to the same basic sound expressed in slightly different ways. In human speech, where the phonemes are known, several different output neurons located close together in a given SOM will recognize variations of the same phoneme. It seems reasonable to assume that a particular meaningful sound found in a nonhuman animal vocalization also would be expressed with some variation.

One area for future research would be to combine the recent progress made in using formants to identify acoustic structures in alarm calls (Slobodchikoff and Placer, 2006) with the knowledge gained from the experiments discussed in this paper. Specific formant information could be combined with the general spectral information utilized by the Kohonen maps to see if an improvement in the performance of the SOMs is achieved. Improvement in SOM performance, reflected by improved classification accuracies, would result in more accurate clustering of like sounds and would provide more detailed information about the small-scale acoustic structure of the vocalizations.

Another fruitful area for future research would be to apply the techniques discussed in this paper to recordings of animal vocalizations that allow the exploration of new types of information encoding. For example, we are now preparing to use Kohonen mapping techniques to study a newly recorded set of alarm calls that will allow us to search not only for predator species information but also for descriptive information about predators and for information about the individuals that are vocalizing. In future work, as the Kohonen methodology is applied to recordings of alarm calls vocalized in a wide variety of situations, it might be possible to develop an acoustic dictionary for prairie dog communications. In such a dictionary each acoustic pattern in a large collection of such patterns would be associated with specific information content. It would then be possible to examine all the patterns found in the dictionary in order to attempt to develop a general theory of the lexical structure of prairie dog communications. The development of this type of rudimentary lexicon for prairie dog communications would also make it possible to attempt a higher level examination of the complete structure of alarm calls to see if a syntactic structure could be determined for these calls. Information about a possible syntactic structure of alarm calls might even allow the synthetic creation of alarm calls for playback experiments.

TABLE 3. Lists of monograms, digrams, and trigrams that are unique to a specific predator species. The list below the bold heading in each cell gives all of the monograms, digrams, and trigrams that begin with the symbol named in the heading and that are unique to the particular predator species also named in the heading of the cell.

S1	S2(Dog) S2S2	S3	S4(Dog) S4	S5	S6	S7	S8
S9	S10(Coyote) S10S10	S11	S12(Coyote) S12S12	S13	S14(Dog) S14S14	S15	S16
S17	S18	S19(Coyote) S19S37	S20	S21(Coyote) S21S21 S21S10	S22(Coyote) S22S22	S23	S24(Coyote) S24S21
S25(Dog) S25S25	S26(Dog) S26S26	S27	S28	S29	S30	S31(Coyote) S31S31	S32
S33	S34	S35(Hawk) S35S35 S35S60 S35S35S60 S35S35S58 S35S35S53 S35S60S53 S35S58S11 S35S53 S35S53S51	S36	S37(Coyote) S37S37 S37S62 S37S19 S37S12	S38(Coyote) S38S38	S39	S40
S41	S42	S43(Dog) S43S43	S44(Dog) S44S63	S45	S46	S47	S48
S49(Dog) S49	S50	S51(Hawk) S51S51 S51S53 S51S51S53 S51S53S53 S51-53-60H	S52	S53(Hawk) S53 S53S60 S53S60S60 S53S60S58 S53S51 S53S51S53	S54(Dog) S54S54	S55	S56
S57	S58(Hawk) S58S35S35 S58S11S32H	S59	S60(Hawk) S60S58 S60S60S58 S60S58S58 S60S58S35 S60S53 S60S53S53 S60S35	S61	S62(Coyote) S62S62 S62S37	S63(Dog) S63S63	S64(Dog) S64S64

Most of the work done by researchers on animal vocalizations has been focused on the macrostructure and general acoustic characteristics of those vocalizations. Consequently, not many tools have been developed with which to study the acoustic microstructure of animal vocalizations. Yet, the vocalizations of a growing number of animals are being shown to contain referential communication which conveys

information about types of predators, degrees of danger, and other details about the world in which these animals live. These discoveries suggest that understanding the small-scale structure of animal vocalizations might be vitally important to determining the level of complexity of those communications and the actual information they contain. New methodologies and software tools for studying the acoustic microstructure of animal vocalizations need to be created so that detailed models and analyses related to these communications can be developed. The experiments described in this paper show how Kohonen maps might well provide one class of these needed tools.

References

- Ackers, S. H., and Slobodchikoff, C. N. (1999). "Communication of stimulus size and shape in alarm calls of Gunnison's prairie dogs," *Ethology* 105, 149–162.
- Baker, M. C., and Logue, D. M. (2003). "Population differentiation in a complex bird sound: A comparison of three bioacoustical analysis procedures," *Ethology* 109, 223–242.
- Blackshaw, J. K., Jones, D. N., and Thomas, F. J. (1996). "Vocal individuality during suckling in the intensively housed domestic pig," *Appl. Anim. Behav. Sci.* 50, 33–41.
- Clark, C. W., Marler, P., and Beeman, K. (1987). "Quantitative analysis of animal vocal phonology: an application to swamp sparrow song," *Ethology* 76, 101–115.
- Eakle, W. L., Mannan, R. W., and Grubb, T. G. (1989). "Identification of individual breeding bald eagles by voice analysis," *J. Wildlife Management* 53, 450–455.
- Galeotti, P., Paladin, M., and Pavan, G. (1993). "Individually distinct hooting in male pygmy owls *Glaucidium passerinum*: a multivariate approach," *Ornis Scandi.* 24, 15–20.
- Galeotti, P., and Pavan, G. (1991). "Individual recognition of male tawny owls (*Strix aluco*) using spectrograms of their territorial calls," *Ethol. Ecol. Evol.* 3, 113–126.
- Hill, F. A. R., and Lill, A. (1998). "Vocalizations of the Christmas Island hawk-owl *Ninox natalis*: Individual variation in advertisement calls," *Emu* 98, 221–226.
- Kangas, J., Torkkol, K., and Kokknen, M. (1992). "Using SOMs as feature extractors for speech recognition," *ICASSP-92* 2, 341–344.
- Kohonen, T. (1988). "The 'neural' phonetic typewriter," *Computer* 21, 11–22.
- Kohonen, T. (1998). "The self-organizing map," *Neurocomputing* 21, 1–6.
- Lengagne, T. (2001). "Temporal stability in the individual features in the calls of eagle owls (*Bubo bubo*)," *Behaviour* 138, 1407–1419.
- Mihelic, F., Ipsic, I., Dobrisek, S., and Pavesic, N. (1992). "Feature representations and classification procedures for Slovene phoneme recognition," *Pattern Recogn. Lett.* 13, 879–891.
- Placer, J., and Slobodchikoff, C. N. (2000). "A fuzzy-neural system for identification of species-specific alarm calls of Gunnison's prairie dogs," *Behav. Processes* 52, 1–9.
- Placer, J., and Slobodchikoff, C. N. (2001). "Developing new metrics for the investigation of animal vocalizations," *Intell. Autom. Soft Comput.* 7, 249–259.
- Placer, J., and Slobodchikoff, C. N. (2004). "A method for identifying sounds used in the classification of alarm calls," *Behav. Processes* 67, 87–98.
- Slobodchikoff, C. N., Ackers, S. H., and Van Ert, M. (1998). "Geographical variation in the alarm calls of Gunnison's prairie dogs," *J. Mammal.* 79, 1265–1272.
- Slobodchikoff, C. N., Kiriazis, J., Fischer, C., and Creef, E. (1991). "Semantic information distinguishing individual predators in the alarm calls of Gunnison's prairie dogs," *Anim. Behav.* 42, 713–719.
- Slobodchikoff, C. N., and Placer, J. (2006). "Acoustic structures in the alarm calls of Gunnison's prairie dogs," *J. Acoust. Soc. Am.* (to be published).

- Sousa-Lima, R. S., Paglia, A. P., and Da Fonseca, G. A. B. (2002). "Signature information and individual recognition in the isolation calls of Amazonian manatees, *Trichechus inunguis* (Mammalia: Serenia)," *Anim. Behav.* 63, 301–310.
- Wu, P., Warwick, K., and Koska, M. (1992). "Neural network feature maps for Chinese phonemes," *Neurocomputing* 4, 109–112.
- Zimmerman, A., and Lerch, C. (1993). "The complex acoustic design of an advertisement call in male mouse lemurs (*Microcebus murinus*) and sources of its variation," *Ethology* 93, 211–224.